DOCUMENT RESUME

ED 076 664 TM 002 672

| | |
|---|---|
| AUTHOR | Mathews, Walter M. |
| TITLE | Computer Narrative Assessment Reports. |
| PUB DATE | Feb 73 |
| NOTE | 13p.; Paper presented at annual meeting of American Educational Research Association (New Orleans, Louisiana, February 25-March 1, 1973) |
| EDRS PRICE | MF-$0.65 HC-$3.29 |
| DESCRIPTORS | Communication (Thought Transfer); *Computer Programs; *Narration; *Reports; *Scores; Speeches; Test Interpretation; *Test Results |
| IDENTIFIERS | California Psychological Inventory; Minnesota Multiphasic Personality Inventory; Preliminary Scholastic Aptitude Test; Teaching Information Processing System; TIPS |

ABSTRACT

The use of narrative test reports overcomes the major barrier to understanding reports, understanding the language that is used. Early attempts to utilize the computer in generating narrative reports include: (1) Teaching Information Processing System (TIPS), involving periodic collection of information from students regarding courses, which is summarized within a few hours into three types of reports--student, section learer, and professor; (2) Preliminary Scholastic Aptitude Test (PSAT) Score Reports, involving eighty distinct sentences, in which variable phrases might be embedded, which are used to compose 75 distinct paragraphs, which in turn are combined to produce the 100 letters needed to interpret all combinations of scores; (3) Programmed Composition of Psychological Test Reports, involving selection of one of eight possible statements for each of the 101 scales of the MMPI (or the 124 of the MMPI and the CPI). Arguments can be made for and against the use of narrative reports. (KM)

FILMED FROM BEST AVAILABLE COPY

ED 076664

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

# COMPUTER NARRATIVE ASSESSMENT REPORTS

Walter M. Mathews

The University of Mississippi

A paper presented at the annual meeting of

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION

February 1973

New Orleans

TM 002 672

When a person takes a test (or in some way is formally assessed), a report is usually produced. The report may be simply the number of correct items on the test, or it may be a detailed description of the testee's responses to a psychological test battery with comments and interpretation from a psychologist. The purpose of assessment reports is communication of information for decision-making, and the communication must be understandable to the receiver of the report.

Assuming no difficulties with the validity and reliability of the test, and a general understanding of the purpose and method of the test, the major barrier to understanding a performance report is understanding the language that is used. The language is usually numerical values of quantitative concepts such as raw scores, grade-equivalent scores, stanines, percentiles and standard errors of measurement. Certainly the target audience of the reports (students or parents or instructors or counselors,etc.) determines the kind of concern that language problems elicit; but to err by assuming an undistorted transmission of information is easy and a common occurrence.

The result of an understandable report should be the transfer of information on performance that can be used in the decision-making process. In the case of standardized achievement tests in education, Goslin found (1967, p. 32) that elementary teachers receive score reports on pupil performance about 80 per cent of the time and have free access to these scores in virtually every case,

yet this information is shared routinely with pupils and their parents less than eight per cent of the time. Current test reporting practices, then, seem to indicate that at least some of the information that could be useful in making educational decisions about a child is not routinely shared with the child and his parents. Proceeding on the assumption that testing information should be shared with parents and children, the obvious prohibiting factor is that "They would not understand the report." That is probably true. After all, some of our teachers and administrators have difficulty understanding current test reports. An alternative to depriving parents and students of testing information is to produce a testing report that is understandable to them. The suggestion and theme of this symposium is the use of reports that are in a narrative format--words that blend to form sentences which join into paragraphs.

## EARLY PROJECTS

Let's look at a few early attempts to utilize the computer in generating narrative reports.

### Teaching Information Processing System (TIPS)

Kelley (1968), a professor of economics at the University of Wisconsin, developed a program, which he calls TIPS, to assist him in teaching. TIPS involves periodic collection of information from students regarding either their understanding of course materials or their reaction to various aspects of course presentations. TIPS provides a means of efficiently utilizing this information for instructional purposes. The information, which is

collected on specialized forms suitable for machine processing, is composed of student responses to a series of multiple-choice questions. Surveys of six to twelve questions take about five to ten minutes to administer. Within a few hours this information is processed and summarized in three separate reports: one for each student, one for each section leader, and a third for the professor.

The student report contains a summary of his performance: his response to each question, the correct answers, and the total number of his correct answers. On the basis of this information, assignments for the forthcoming period are also indicated. The assignments (some required--some optional) vary considerably in nature, level, and intensity. A student scoring well may receive optional assignments and/or required work at a higher level. The student performing poorly may receive not only a heavy dose of required work but also a set of materials designed to bring him toward the mean class performance.

Additional information on the student report is generated on the basis of past as well as current performance. If the student has performed poorly over several surveys, he will be instructed on the student report to establish an appointment with the instructor or teaching assistant. If the student performed consistently and exceptionally well, he may be notified that a short paper may be substituted, at his option, for the midterm examination. A sample of a student report is presented as Figure 1.

---

Figure 1 about here

---

The teaching assistant report contains information to help him appraise the performance of his individual sections, including statistics on percentage correct by question or by concept, actual responses on the survey, lists of students required to establish appointments or tutorials, and so forth.

The professor's report is similar to that received by the teaching assistant, although the information available applies to all students enrolled in the course rather than only to the students in particular sections. With this information the professor may elect to alter lectures, section coverage, problem sets, or other teaching instruments for the forthcoming period. In summary, TIPS is a system for gathering and reporting objective and timely information useful for more effective teaching.

Preliminary Scholastic Aptitude Test (PSAT) Score Reports

The effective reporting of test results to admissions officers, guidance counselors, and the individuals who take educational tests is a rather more complex matter than might appear at first glance. It is apparent that reporting only numerical scores is hardly adequate. The statistical, psychological, and educational contexts which allow the user to infer relevant meanings must be provided as well. The problem is particularly acute for the programs in which the primary reporting target is the test taker himself.

Since the precise interpretation of mental test scores does require rather sophisticated insights into statistics, psychology, and education, it is common practice to have test scores reported

to individual test takers by guidance counselors who are expected to have the necessary sophistication. Unfortunately, not all guidance counselors do have sufficient psychometric sophistication, nor do they have the time to prepare detailed analyses and to give individual interpretation of mental test results. Of course, most educational testing services prepare a wide range of interpretive materials for their testing programs to aid both the counselor and the student in interpreting test scores. But even with these aids, some statistical sophistication is still needed for adequate understanding.

Helm and Harasymiw (1968) designed variable format computer-generated letters to be sent to examinees of the Preliminary Scholastic Aptitude Test as a report of their performance on the test. They found it necessary to prepare eighty distinct sentences where a sentence might have variable phrases imbedded. These eighty sentences were used to compose seventy-five distinct paragraphs which, in turn, were combined to produce the 100 letters needed to interpret all combinations of verbal and mathematical scores. They wrote a computer program to generate these letters and a sample of their report appears as Figure 2.

---

Figure 2 about here

---

Programmed Composition of Psychological Test Reports

A computer-generated verbal diagnostic report on a standardized psychological test has been used routinely at the Mayo Clinic in

Rochester, Minnesota since 1962 (Swenson, 1962), and is available commercially from Behaviordyne, Inc. The Mayo program utilizes the Minnesota Multi-phasic Personality Inventory (MMPI) which is an objective pencil-paper psychological test. The machine-produced report is a group of disconnected statements, or decisions, about the subject as measured by the scales of the MMPI. Finney improved the Mayo program by offering an alternative of adding scales from the California Psychological Inventory (CPI), and improving the coherence of the report (1966). A large number of scales are scored, 101 with the MMPI alone or 124 with the MMPI and the CPI.

The report is built by selecting statements and then combining them into paragraphs. For each of the 101 scales, one statement is chosen from among eight possible statements, depending on the individual's score on the scale. By this method, 101 statements are chosen from a repertory of 808. Finney's program has the computer compose a full report on each individual's personality--the kind of report that a psychologist might write after seeing a person several times and administering a full battery of tests. The first one-third of a sample report is given as Figure 3.

---

Figure 3 about here

---

Finney and Auvenshire have subsequently developed several different kinds of reports written for different purposes. They are now extending their work to other objective psychological tests.

Baker suggested (1971) that electronic computer technology be utilized to generate reports on other types of testing instruments in order to make the results more meaningful to the persons examined and to facilitate better use of the results of testing. He concluded:

> The mechanics of having the computer program prepare verbal descriptions depends upon several factors. First, the insight of the test constructor into the area of interest; second, the relation of levels of test and diagnostic scores to pupil performance; third, the cleverness of the computer programmer in generating connecting prose from somewhat disconnected verbal descriptions.

Some Potential Advantages

Depending upon the specific application, several advantages accrue from the use of narrative reports. A few general advantages follow:

*Clear Communication--The report is in understandable English and does not ask the recipient to search the page for clues on where to begin "reading" the report.

*Personal--The use of the testee's name and the appropriate personal pronouns can easily be included in the text to humanize the report and increase the attention paid to it. It does not "look like a computer report."

*Efficient--A more complete report is available at the cost of fewer professional-person-hours. Narrative reports are not suggested as a substitute for the professional, but rather as an aide to the professional.

*Self-Explanatory--The report can be reasonably self-contained with less need for individual guidance to assist in its initial interpretation.

*Flexible--The underlying philosophy of the narrative can be varied as needed for different audiences and audience sub-sets.

*Public Relations--An inherent public relations value can accrue to the organization from a professional report that is understandable and personalized.

Certainly there are opposing arguments that could be offered:

*Muddled Communication--The report has to be read, thereby restricting its audience. Realizing that, the reading level of the report still makes communication less than efficient with another segment of people. Others who can read well, do not want to be bothered to read--they just want the score, and without extraneous verbiage.

*Too Personal--Some people may react negatively to the thought of a machine attempting to be "intimate" with them.

*Expensive--Initial development of the system is expensive with no promise of a lowered cost per report.

This list could also be extended, and perhaps you will add to it during the discussion after the papers. We do not pretend to propose a panacea for reporting problems. We do, however, think that narrative reports have a place, and that their application should be furthur explored to determine its precise location.

Figure 1

SAMPLE COPY OF KELLEY'S TEACHING INFORMATION PROCESSING SYSTEM (TIPS) STUDENT REPORT

TIPS
STUDENT PERFORMANCE SURVEY RESULTS
PRINCIPLES OF ECONOMICS (103)
PROFESSOR ALLEN C. KELLEY

TURNER, LAWRENCE
715648
SURVEY TAKEN 10/25/67
SECTION NUMBER AND TIME 2, 9:55 F
SECTION LEADER MISS GREEN

OUT OF A TOTAL OF 10 QUESTIONS, YOU CORRECTLY ANSWERED 3. THE QUESTION NUMBER, YOUR RESPONSE AND THE CORRECT ANSWER ARE PROVIDED IN THE TABLE BELOW. YOU ARE URGED TO MAKE SURE YOU UNDERSTAND THE NATURE OF ANY INCORRECT RESPONSES YOU MAY HAVE MADE.

SUMMARY OF SURVEY RESULTS

| QUES. NUM. | YOUR ANSW. | CORR. ANSW. | QUES. NUM. | YOUR ANSW. | CORR. ANSW. |
|---|---|---|---|---|---|
| 1 | G | G | 6 | B | A |
| 2 | C | B | 7 | D | E |
| 3 | D | B | 8 | A | A |
| 4 | A | C | 9 | D | C |
| 5 | F | F | 10 | C | B |

YOUR ASSIGNMENT FOR THE WEEK, TO BE HANDED IN DURING THE DISCUSSION SECTION ON 11/03/67, IS THE FOLLOWING --

PROBLEMS 1, 3 AND 4 ON HANDOUT 2C

ADDITIONALLY, YOU ARE REQUIRED TO WORK THROUGH CHAPTER 2 OF - MICRO-ECONOMICS, A PROGRAMMED BOOK, BY LUMSDEN, ATTIYEH, AND BACH. IT WOULD BE USEFUL TO CONSULT THE PROGRAMMED UNIT BEFORE YOU READ HANDOUT 2C.

TO TEST YOUR UNDERSTANDING OF THE MATERIALS, YOU MAY, AT YOUR OPTION, ELECT TO COMPLETE THE FOLLOWING PROBLEMS --

1. WORKBOOK, PP. 37-38.
2. HANDOUT 2A, PROBLEMS 3 AND 4.

NOTICE -- AN ALL UNIVERSITY LECTURE BY PROFESSOR MILTON FRIEDMAN, PAST PRESIDENT OF THE AMERICAN ECONOMIC ASSOCIATION, WILL BE HELD IN 6210 SOCIAL SCIENCE, 7:30 P.M., 11/01/67. THE TOPIC -- MONETARY AND FISCAL POLICY RECONSIDERED.

Figure 2

SAMPLE COPY OF THE PRELIMINARY SCHOLASTIC
APTITUDE TEST (PSAT) SCORE REPORT

E D U C A T I O N A L  T E S T I N G  S E R V I C E

P R I N C E T O N, N . J .  0 8 5 4 0

JANUARY 21, 1966

DEAR MR. LETR 99,

WE WANT TO REPORT TO YOU THE SCORES YOU EARNED ON THE PRELIMINARY SCHOLASTIC APTITUDE TEST YOU TOOK ON OCTOBER 9, 1965. YOUR APTITUDE FOR COLLEGE WORK IS OUTSTANDING. IF YOUR HIGH SCHOOL MARKS ARE CONSISTENT WITH THE HIGH SCORES YOU HAVE EARNED ON THE TEST YOU WILL HAVE LITTLE DIFFICULTY IN BEING ACCEPTED AT A COLLEGE OF YOUR CHOICE.

YOU EARNED A SCORE OF 68 ON THE VERBAL SECTION OF THE TEST. THE VERBAL SECTION OF THE TEST MEASURES YOUR ABILITY TO READ WITH UNDERSTANDING AND TO USE WORDS EFFECTIVELY. A SCORE AS HIGH OR HIGHER THAN THE ONE YOU HAVE EARNED IS EARNED BY LESS THAN 11 PER CENT OF JUNIORS OF YOUR SEX WHO LATER ENTER COLLEGE. VERBAL APTITUDE IS PARTICULARLY IMPORTANT FOR SUCCESSFUL COLLEGE WORK IN THE HUMANITIES AND FINE ARTS.

YOU EARNED A SCORE OF 68 ON THE MATHEMATICAL SECTION OF THE TEST. THE MATHEMATICAL SECTION OF THE TEST MEASURES YOUR ABILITY TO REASON AND WORK EFFECTIVELY WITH NUMBERS. A SCORE AS HIGH OR HIGHER THAN THE ONE YOU HAVE EARNED IS EARNED BY LESS THAN 11 PER CENT OF JUNIORS OF YOUR SEX WHO LATER ENTER COLLEGE. MATHEMATICAL APTITUDE IS PARTICULARLY IMPORTANT FOR SUCCESSFUL COLLEGE WORK IN THE SCIENCES AND ENGINEERING.

YOU SHOULD NOT THINK OF YOUR TEST AS EXACT POINTS BUT AS A RANGE OF SCORES EXTENDING ABOUT THREE POINTS ABOVE AND THREE POINTS BELOW THE SCORE WE HAVE REPORTED TO YOU. THEY GIVE A GOOD INDICATION OF HOW YOU MAY EXPECT TO SCORE ON THE SCHOLASTIC APTITUDE TEST. THE CHANCES ARE FOUR OUT OF FIVE THAT YOU WILL SCORE BETWEEN 630 AND 730 ON THE VERBAL SECTION OF THE SCHOLASTIC APTITUDE TEST WHEN YOU TAKE IT NEXT YEAR. THE CHANCES ARE FOUR OUT OF FIVE THAT YOU WILL SCORE BETWEEN 630 AND 730 ON THE MATHEMATICAL SECTION OF THE SCHOLASTIC APTITUDE TEST WHEN YOU TAKE IT NEXT YEAR. INSOFAR AS VERBAL AND MATHEMATICAL APTITUDES ARE CONCERNED YOU CAN HAVE CONFIDENCE IN YOUR ABILITY TO DO SUCCESSFUL COLLEGE WORK.

THE SCORES YOU HAVE EARNED SHOULD ENCOURAGE YOU TO APPLY FOR ADMISSION TO AN OUTSTANDING COLLEGE. YOUR FUTURE EDUCATIONAL PLANS SHOULD CONSIDER ADVANCED GRADUATE WORK. AFTER YOU HAVE DISCUSSED YOUR SCORES WITH YOUR PARENTS AND YOUR COUNSELOR OR PRINCIPAL SHOULD YOU HAVE ADDITIONAL QUESTIONS ABOUT THEM YOU MAY WRITE TO EDUCATIONAL TESTING SERVICE, PRINCETON, N. J.

SINCERELY YOURS,

E. T. S.

Figure 3

## SAMPLE COPY OF FINNEY'S REPORT ON THE MINNESOTA MULTI-PHASIC PERSONALITY INVENTORY

This is a report of MMPI and CPI testing of a female age 16, case number B-000005. This test, like any test, is subject to error. Testing only supplements other diagnostic examinations.

First let us examine the evidence of validity and the attitude with which she took the test.

*RA-0, Cm-4*
*Gbs-10*
*L-3*
*Gi-4*
*Wb-3*

On the CPI she gives mostly the common and conventional answers. That may be a sign of at least average common sense and judgment, and of being sufficiently steady, reliable, and realistic. She does not give a consistently favorable nor a consistently unfavorable picture of herself. She is afraid to admit even small flaws in herself, in terms of standards which are naive, rigid, perfectionistic, moralistic, unrealistic, and overly conventional. That shows a lack of insight. It also shows that she does not distinguish clearly between fundamental obligations, which people can and will expect her to meet, and the lesser or shallower matters in which a falling short in performance is tolerable. She wants to make a good impression in taking the test, and she gives the impression of having at least an average degree of warmth. She is moderately ambitious, alert, and productive, and likes working. She has no serious doubts about herself.

*A-5, 7-4°*
*R-4, Rbw-5°*
*B-4°, Cfi-5*
*3f-27*

She does not tell of anxiety or stress and is not looking for help. She is a reasonably compliant person. She has a normal amount of flexibility.

In terms of these factors, she seems to be a normal, average, flexible person.

*2pc-31*

But the two-point code tells us as follows. She has hysterical conversion reactions of some specific location or other. She is naive, exhibitionistic, self-centered, and demanding, and tends to manipulate and exploit people. Because of repression she lacks insight and is not motivated for psychotherapy.

Now, what is the evidence for psychosis or mental illness?

*Rec*
*Scpt-3*
*Sx-3*
*43-7*

None of the measures indicate that she is psychotic. Some measures are doubtful, as follows. The obsessive and schizophrenic indicators are about at equal level. On a schizophrenic correction scale, she seems to have psychotic trends. But most measures indicate that she is not psychotic. She does not use rituals or compulsive acts at all to ward off anxiety.

Next we consider narcissism, guilt, and basic trust.

*N-7*
*Sa-5*
*29-6*
*Gu-6*
*27-4*
*Em-7, 4-5°*
*Dom-4, Es-5*
*6-3*
*PtK-4°*
*44-4°, Ob-6*
*42-3, Ph-6*

Her self-esteem is low and she doesn't feel proud of herself. But by another measure her self-acceptance is within the average range, though she tends to blame herself a little more than the average. She shows signs of less than average guilt feeling. And she tends to deny guilt. She has a normal amount of concern with what people think of her. But she denies any feeling of self-consciousness or embarrassment. She has at least average dominance and initiative. She scores almost average on ego strength, and has fair tolerance for frustration. This is a good level of ego strength for a psychiatric patient. She has the assets of benefit from psychotherapy, but only if motivation and distress are also present. To a moderate degree she maintains an optimistic attitude by denying discouragement. She tells of very little worrying; less than the average person. She shows signs of having some fears or phobias. But she does not admit fears or phobias.

Now we turn our attention to problems of dependency.

*1-4*
*2-4*
*Rfd-5°, 3-4°*
*Ud-5°, Ody-4°*

The signs are that she has only slightly more dependency need than the average, if at all so. Within the average or normal range she seems to put her dependency needs into action.

Now, what about being demanding or orally aggressive?

## REFERENCES

Baker, Frank B., "Automation of Test Scoring, Reporting and Analysis." Chapter 8 in Educational Measurement, 2nd ed., R.L. Thorndike ed. (Washington, D.C.: American Council on Education, 1971).

Finney, Joseph C., "A Programmed Interpretation of the MMPI and the CPI," Archives of General Psychiatry, XV (1966), pp. 202-36.

Goslin, David A., Teachers and Testing. (New York: Russell Sage Foundation, 1967).

Helm, Carl E. and Harasymiw, Stefan J., "Computer-based Score Reports," Measurement and Evaluation in Guidance, I, 1 (Spring, 1968), pp. 27-35.

Kelley, Allen C., "An Experiment with TIPS: A Computer-aided Instructional System for Undergraduate Education," The American Economic Review: Papers and Proceedings, LXXX (1968), pp. 446-57.

Swenson, W.M., et al., "Symposium on Automation Techniques in Personality Assessment," Proceedings of Staff Meetings of the Mayo Clinic, XXXVII (1962), pp. 61-82.